

Katalin Balog

Disillusioned

Abstract: *In ‘The Meta-Problem of Consciousness’, David Chalmers draws a new framework in which to consider the mind–body problem. In addition to trying to solve the hard problem of consciousness — the problem of why and how brain processes give rise to conscious experience — he thinks that philosophy, psychology, neuroscience, and the other cognitive sciences should also pursue a solution to what he calls the ‘meta-problem’ of consciousness — i.e. the problem of why we think there is a problem with consciousness. My claim is that, while Chalmers’ project is generously ecumenical as well as beautiful in its meticulous detail, it is mistaken in its core assumption that the meta-problem can be formulated as an ‘easy problem’ for science to solve. Furthermore, the project tilts the field toward illusionism against type-B materialism, as far as physicalist solutions to the hard problem and the meta-problem are concerned. I will argue that type-B materialism emerges unscathed from this dialectic.*

1. The Meta-Problem

Puzzlement about conscious experience and the intuition that there is something anomalous in its relationship to the physical world appears to be universal. It takes many forms, but its clearest, best articulated manifestations are related to the well-known epistemic gaps between the physical and the phenomenal — the special knowledge experience affords those who have it that no amount of physical knowledge can provide, the conceivability of the absence of any experience in otherwise physically identical beings, the inexplicability of conscious experience in terms of the underlying brain processes, as well as the seeming lack of a physical explanation of how we can be aware of our experiences *directly, incorrigibly, and in a way that reveals their*

Correspondence:
Email: kbalog@rutgers.edu

Journal of Consciousness Studies, 27, No. 5–6, 2020, pp. 38–53

essence. This is well-travelled territory for any philosopher who has pondered the mind–body problem.

So far so good. It is sensible to require that a theory of consciousness should explain (or explain away) these ‘problem intuitions’, as Chalmers calls them, as well as consciousness itself: any solution to the problem of consciousness needs to be compatible with — and plausibly closely integrated with — a solution to the meta-problem. Everyone concerned about the mind–body problem could — and should — sign up for the programme thus described. It turns out, however, that Chalmers proposes a special constraint on the meta-problem research programme: he thinks that the problem intuitions, as well as any proposed explanations for them, should be stated in *topic-neutral* language, i.e. without using the term ‘conscious experience’, ‘phenomenality’, or any of their cognates.

The reason for this requirement is that he wishes the meta-problem research programme to be ‘for everyone’, and so he wishes the formulation of the problem to be neutral between realism about consciousness and the increasingly popular eliminativist view (which goes under the slightly more innocuous sounding name ‘illusionism’ these days). Posing the problem in this way, the thinking goes, leaves room for both realist and illusionist views because it neither presupposes the existence of consciousness from the start, nor does it require that consciousness is causally or explanatorily irrelevant to the meta-problem intuitions. A further advantage is that this formulation turns the meta-problem — in contrast to the hard problem — into an easy problem: a problem that can be formulated in terms of function or structure, so its solution can proceed along straightforward scientific paths. All this seems sensible: we might as well work on a problem that is scientifically tractable and that all sides of the debate can accept as their own.

However, I think that, while this research programme is likely to generate interesting results, it is unlikely to live up to its promise. In the next section, I will argue that the problem intuitions, even if they can be given topic-neutral formulations, could not be fully explained topic-neutrally. I argue that the meta-problem doesn’t have a scientific solution that doesn’t appeal to consciousness or related concepts. In the third section, I defend physicalism in the face of this claim; that is, I outline how physicalism is compatible with there being no scientific solution to the meta-problem. In the last section, I comment on illusionism and the quest for a scientific understanding of the mind.

2. Problem Intuitions Do Not Have a Topic-Neutral Counterpart

Here is Chalmers' suggestion about how to produce the requisite topic-neutral versions of the problem intuitions:

One could put phenomenal intuitions in an existential form, such as 'We have special properties that are hard to explain' or 'that are non-physical', 'that provide special first-person knowledge', 'that could be missing in robots', and so on. Alternatively, one could simply require that phenomenal intuitions be explained up to but not including the fact that they are specifically about consciousness. Once we have explained judgments of the form 'We have special first-person knowledge of X which is hard to explain in physical terms', and so on, we have done enough to solve the meta-problem... [W]e can call these quasi-phenomenal judgments. Quasi-phenomenal judgments do not so obviously depend on consciousness, and might even be shared by zombies. (Chalmers, 2018, p. 18)

Notice that these are not beliefs many of us would ever entertain explicitly. Nevertheless, they follow from the usual problem intuitions, so presumably it would be easy to get people to assent to them, having pondered the original (non-topic-neutral) problem intuitions. And, at first sight, it looks reasonable to expect that these 'quasi-phenomenal' problem intuitions can be explained topic-neutrally in terms of perception and introspection. Chalmers summarizes the most promising topic-neutral account to figure in such an explanation as follows:

We have introspective models deploying introspective concepts of our internal states that are largely independent of our physical concepts. These concepts are introspectively opaque, not revealing any of the underlying physical or computational mechanisms. Our perceptual models perceptually attribute primitive perceptual qualities to the world, and our introspective models attribute primitive mental relations to those qualities. We seem to have immediate knowledge that we stand in these primitive mental relations to primitive qualities, and we have the sense of being acquainted with them. (*ibid.*, p. 34)

What stands out in this description is the expression 'we have the sense of being acquainted' with our experiences. We think of acquaintance as the relationship we have to our experience that is both direct and intimate and seems to reveal its essence in a special way. The experience seems present in the very awareness of it. It is hard to think of a way in which it can be characterized topic-neutrally. In giving his gloss of the best contenders for topic-neutral solutions to

the meta-problem, Chalmers himself characterizes acquaintance in what seems to be phenomenal terms:

There are a couple of distinct elements to the sense of acquaintance. One is the sense of *presentation*: that we are somehow immediately presented with our experiences. Another is the sense of *revelation*: that the full nature of consciousness, and of various phenomenal properties, is fully revealed to us in introspection. (*ibid.*, p. 25)

While such an account of acquaintance is technically topic-neutral in that it doesn't explicitly involve phenomenal concepts, if topic-neutrality requires that we only use functional or structural concepts that are apt to provide appropriate links to the special sciences so as to generate a scientific account of the phenomenon in question — as solution of the 'easy' meta-problem seems to require — then this account is not truly topic-neutral. These concepts seem essentially phenomenal and not functional.

This is a serious problem for Chalmers' project. Even if we manage to purge them from the problem intuitions, the concept 'acquaintance', 'revelation', 'presence', or the like, as I will now argue, needs to figure in any full *explanation* of the problem intuitions, topic-neutral or otherwise.¹ The reason is that any explanation of the problem intuitions needs to be specific to the consciousness case. These intuitions arise *only* with regard to consciousness, but not with regard to any other mental state that shares characteristics with consciousness that can be articulated and explained in topic-neutral terms, while lacking phenomenality. I am going to present two such cases. If such cases are possible, it follows that an explanation that doesn't appeal to consciousness or a related concept such as acquaintance will explain 'too much'. It will predict the problem intuitions to arise even in cases where they in fact don't.

Here is an illustration of the point. Say someone likes all carnations except bright red ones which they intensely abhor. This person, whenever offered a red carnation, has the disposition to form the 'problem intuition' 'I don't want this red carnation in my house'. Let's suppose you, for whatever reason, want to study a colour-neutral version of this problem intuition, say 'I don't want this carnation in my house'. Our red carnation detractor also has a disposition to form this problem intuition as well. In both cases, it seems the explanation will have to

¹ The above quote from p. 34 seems to indicate Chalmers is aware of this.

appeal to how perceptual features of carnations affect you. But any such explanation better appeal to not just how certain general colour-neutral perceptual features of carnations affect you. That could not be a full explanation of the tendency to endorse the colour-neutral problem intuition as all those features, when instantiated by a, say, pink carnation, would not elicit the problem intuition. The explanation, *even of the colour-neutral problem intuition* entertained by our red carnation detractor, better appeal, *inter alia*, to the very redness of the carnation.

Back to the consciousness case: even if one drops ‘acquaintance’ and its cognates from the formulation of the problem intuitions, their full explanation has to be specific to the consciousness case. It better not predict such problem intuitions in other cases that share all relevant topic-neutral descriptions with conscious experience except the very consciousness and acquaintance itself.

Take a problem Chalmers discusses just prior to giving his account of acquaintance: the problem of why introspecting thoughts doesn’t create the same problem intuitions as introspecting experience. He thinks it is because the one acquaints us with its subject, and the other doesn’t. But that would only be compatible with the meta-problem research programme if something could be found in those otherwise similar, topic-neutrally described cognitive mechanisms involved in thinking about experience and thought respectively — both involving direct concepts applied to mental states as they occur — that explained the fact that the one acquaints us with its object and the other doesn’t. It appears clear, however, as we discussed earlier, that any formulation in strictly functional/structural language will not capture ‘acquaintance’, ‘revelation’, or ‘presence’. It is not possible to purge the explanation of the problem intuitions of its phenomenal character.²

But one might say that thought is necessarily phenomenal after all and, relatedly, argue that the problem intuitions should and do — on

² Chalmers considers this response to his meta-problem research programme. As he puts it, ‘Some non-reductionists may embrace meta-problem nihilism: there is no solution to the meta-problem. Alternatively, if we understand the meta-problem more broadly as “Explain our problem intuitions in topic-neutral terms, or explain why this is impossible”, then the meta-problem nihilist says that any solution must take the second horn. As discussed earlier, a version of this view might be taken by anomalous dualists and anomalous materialists for whom behaviour and/or the causal role of consciousness cannot be systematized in topic-neutral terms. It is far from clear how this would work, but there is at least room to investigate the possibility’ (Chalmers, 2018, p. 41).

proper reflection — arise with regard to thought as well. If this is the case, thought would not constitute the contrasting case for the explanation of problem intuitions that I need to make my case.

There is another example illustrating the same point, however. As I argue in Balog (1999), there could be special concepts of brain states (I call them ‘yogi-concepts’) that, described topic-neutrally, appear to be just like phenomenal concepts. They pick out brain states that they refer to directly, in the act of having of those states, as it were. The referent is, like in the phenomenal case, *present* in the thought, as much as the concept ‘presence’ can be made sense of in a topic-neutral fashion. The difference is, the states they pick out are not phenomenal states, and, as a matter of fact, yogis admit that they have no idea about the precise nature of these brain states. These concepts work somewhat like blindsight concepts, except that they refer to brain states. Though we do not have such concepts, it doesn’t seem impossible for thinkers to have such concepts. And they seem to fit the outlines of the explanatory account Chalmers deems most promising (2018, p. 34). I have left out references to perception as opposed to introspection to fit the case at hand:

We have introspective models deploying introspective concepts of our internal states that are largely independent of our physical concepts. These concepts are introspectively opaque, not revealing any of the underlying physical or computational mechanisms. Our introspective models attribute primitive mental relations to the qualities introspected. We seem to have immediate knowledge that we stand in these primitive mental relations to the qualities introspected.

It seems like the functional profile of yogi concepts matches the functional profile of phenomenal concepts almost exactly — except for the fact that thinking about experience gives rise to problem intuitions, whereas thinking about yogi-concepts arguably does not. To explain the difference, one needs to appeal to phenomenality, acquaintance, presence, or revelation, etc. — no topic-neutral conception will do. However, it is exactly this, the necessity to enlist the sense of acquaintance as an explanans, that dims the promise of the meta-problem research programme. If we can’t get the explanation right without appealing to acquaintance, we have given up on the meta-problem research programme as an easy problem for science to work on and solve.

This should not be very surprising, given that we do not entertain quasi-phenomenal judgments in isolation from the problem intuitions themselves. We only make quasi-phenomenal judgments, if at all, on

the basis of the problem intuitions. So the order of explanation should go through the problem intuitions themselves; and not even Chalmers claims that the problem intuitions, couched as they are in phenomenal language, could possibly be explained in topic-neutral terms.

Chalmers himself criticizes various reductive accounts that take up the ‘meta-problem challenge’ on the grounds that, without appealing to acquaintance and primitive property attribution, explanations of the problem intuitions are unlikely to work. He points out that, for example, the integrated information theory (Tononi, 2007), global workspace theories (Baars, 1988), or higher-order thought theories (Rosenthal, 2002) all fail to provide adequate explanations of the problem intuitions. But these theories are not in the business to do that. Explaining the problem intuitions topic-neutrally runs into the same issue that explaining consciousness itself in topic-neutral terms does: it runs into the explanatory gap. The meta-problem cannot be solved scientifically, for the very same reasons that the hard problem itself cannot be solved scientifically. The problem intuitions (in their original, unpurged form) arise on the basis of thinking about consciousness in phenomenal terms. Therefore the problem intuitions cannot be explained in purely physical/functional terms any more than consciousness itself can be explained in purely physical/functional terms. In both cases, this is due to the conceptual isolation of phenomenal concepts from physical/functional concepts.

There is a possible defence that might be adduced in favour of the meta-problem research programme.³ Conceding the point about explaining the problem intuitions *qua* problem intuitions, one might nevertheless suggest that they could be formulated in strictly neuroscientific or physical-behavioural terms. That will make it unproblematic that they can be explained in topic-neutral terms. This is an interesting proposition; I don’t think, however, that it can accomplish what it sets out to.

I agree that, at least in principle, it is possible to provide a physical/functional description of the problem intuitions. For physicalists this is self-evident. Non-physicalists should also concur that it is possible, at least in principle, to provide a physical/functional description of the brain that is nomologically correlated with entertaining the problem intuitions.⁴ In either case, it is fairly

³ Thanks to François Kammerer for raising this point.

⁴ Interactionist dualists wouldn’t agree with this but I will leave this point aside now.

uncontroversial that it is possible, at least in principle, to fully explain the problem intuitions or their neural correlates so stated in physical/functional terms. This seems to count in favour of the meta-problem research programme: we know that the problem intuitions or their neural correlates so stated have a topic-neutral explanation, so it is a reasonable goal to set out to find it. I don't think, however, that this move succeeds in circumventing my criticism of the meta-problem research programme.

We are nowhere near knowing enough about the brain to identify any thought, including the problem intuitions, in physical/functional terms, and it is highly questionable when if ever we will be able to do so. This puts the research programme in the (perhaps distant) future. But, and this is the crucial point, even if we had a full mapping of thought to neural functioning, the explanatory target, *formulated in these terms*, would have little to do with the explanatory target of the meta-problem research programme, which is quasi-phenomenal judgments, *described in common sense language*. It is in this form that the meta-problem research programme appears relevant to a solution of the hard problem. Explanation is a concept-dependent affair. Chalmers' research programme is about understanding, in functional terms, the cognitive mechanisms that explain why we so naturally come to entertain the quasi-phenomenal judgments. On the other hand, the explanatory target in this proposed new guise is not belief, or (meaningful) speech, but neurophysiological processes. That is one (or a few) levels down from the original explanatory target. They are not the same project. So this way of rescuing the meta-problem research programme certainly undercuts its spirit if not its letter.

An upshot is that leaving the language of consciousness, acquaintance, etc. behind in this way won't allow a proponent of illusionism to argue that consciousness or acquaintance does not figure in the explanation of the target phenomena. They are not implicated *as such* when we describe the brain in neural terms, but that doesn't mean they are not implicated when the description is appropriately higher level.

In the next section, I continue the discussion of physical/functional descriptions of consciousness, acquaintance, and related phenomena in the context of the anti-physicalist arguments.

3. The Phenomenal Concept Strategy and the Meta-Problem

While Chalmers has characterized the meta-problem research programme so as to include illusionism, he gives short shrift to what I think is the strongest, most plausible physicalist account, i.e. the view he used to call type-B physicalism (Chalmers refers to it in this paper as reductive physicalist realism). In fact, the most plausible version of such a view, based on the phenomenal concept strategy,⁵ has built an explanation of the problem intuitions into its very account of conscious experience. Though there is no space here to go into the details, I want to briefly indicate how the phenomenal concept strategy can be defended from Chalmers' criticism. Here is Chalmers' summary:

The well-known phenomenal concept strategy tries to explain many of our problem intuitions in terms of features of our phenomenal concepts. If this works, and if the relevant features can then be explained in topic-neutral terms, we will then have a solution to the meta-problem. I have criticized the phenomenal concept strategy elsewhere (Chalmers, 2007), arguing that there are no features of phenomenal concepts that can both be explained in physical terms and that can explain our epistemic situation when it comes to consciousness... In [that] paper, I argued on the most common accounts where the features of phenomenal concepts can be physically explained, the concepts are too 'thin' to explain our problem intuitions. For example, the suggestion that phenomenal concepts are indexical concepts such as 'this state' does not really explain our knowledge intuitions and others: when we pick out a state indexically as 'this state', we are silent on its nature and there is no obvious reason why it should generate problem intuitions... Something similar goes for many extant suggestions. (2018, p. 21)

Chalmers' criticism, laid out in detail in his earlier (2007) paper, is based on what appears to be an uncharitable understanding of the phenomenal concept strategy. A proponent of the phenomenal concept strategy has a ready response⁶ to the charge that there are no features of phenomenal concepts that can both be explained in physical terms and that can explain our epistemic situation when it comes to consciousness. The answer is based on the idea that the explanations in question depend on *how we are describing* phenomenal concepts, as well as our epistemic situation.

⁵ Loar (1990) originated this approach. Various people developed versions of the view, see, for example, Block (2006), Chalmers (2003), Papineau (2002), and Balog (2012a).

⁶ See e.g. Balog (2012b).

Yes, it is true that phenomenal concepts, *when thought about in terms of phenomenal concepts*, are not physically explicable. Neither can our epistemic situation, *when thought about in terms of acquaintance, etc.*, be physically explained. But that is not a problem for physicalism *even though*, as I argued above, we can't leave 'acquaintance' out of our explanations of the problem intuitions without first having solved the hard problem.

That phenomenal concepts, *phenomenally conceptualized*, are not explicable in physical terms is simply a consequence of the explanatory gap between the physical and the phenomenal. The physicalist can explain the explanatory gap itself by appeal to unique features of phenomenal concepts. What makes this perfectly compatible with physicalism is the fact that *when we think about phenomenal concepts in physical terms* — as physicalism claims is possible, at least in principle — those same concepts *are* physically explicable. There is no contradiction here.

Physicalism asserts that experiences are physical processes, and, following the phenomenal concept strategy, it further asserts that tokens of phenomenal concepts are partly constituted by tokens of the experiences they refer to, that is, tokens of these same physical processes. Because of this, the cognitive mechanisms involved in phenomenal concepts are completely isolated from the cognitive mechanisms involved in physical concepts, and so an explanatory gap arises. In addition, this account helps explain acquaintance in physicalist terms. On this view, when we think about experience deploying phenomenal concepts, the token experience and the token concept we use to think about it share something very substantial: the phenomenal character of the experience. Acquaintance is the special, intimate epistemic relation we have to our phenomenal experience through the *shared phenomenality* of experience and thought about experience. Shared phenomenality is what explains the presence and the revelation involved in acquaintance. Again, nothing in this idea is incompatible with the basic assertion of physicalism, i.e. that experiences, with their phenomenality, are physical processes.

This is a robustly realist account of experience that has a ready philosophical solution to the meta-problem, either in its original formulation or in the topic-neutral variety. It also explains why, as I argued in the previous section, the meta-problem is not an easy problem for science to solve. Furthermore, it is the only view that doesn't run afoul on what Chalmers labels the 'coincidence problem'. Here is how he describes it:

...it is hard to avoid a sense of coincidence entirely. As long as we have modal independence, so that the meta-problem processes *could* have come apart from consciousness, it can seem lucky that they have not. Where psychophysical laws are concerned, it seems lucky that the laws are as they are. Only this luck ensures that we are not in a zombie world with physical processes and phenomenal intuitions but no consciousness, or in an inverted world where these processes yield pleasure when we feel pain. Where realization is concerned, it seems lucky that the meta-problem processes are in fact realized by consciousness rather than by something else. (2018, p. 48)

This problem, alongside the problem of mental causation, forms the basis of the strongest arguments against dualism. Type-B physicalism completely avoids this issue while also rejecting the wildly implausible claims of illusionism. So it is surprising when Chalmers says: ‘I think the most important views here are realizationism (for the non-reductionist) and strong illusionism (for the reductionist).’⁷

4. Illusionism and the Meta-Problem

Chalmers thinks illusionists typically hold their view because of their belief in the soundness of a form of debunking argument, based on the premise that ‘[t]here is a correct explanation of our beliefs about consciousness that is independent of consciousness’ (2018, p. 45), i.e. that the meta-problem has a topic-neutral solution. The argument, in a schematic form, goes like this:

1. There is a correct explanation of our beliefs about consciousness that is independent of consciousness.
 2. If there is a correct explanation of our beliefs about consciousness that is independent of consciousness, those beliefs are not justified.
-
3. Our beliefs about consciousness are not justified.

⁷ It is also a bit surprising when Chalmers calls what he now calls reductive realist physicalism — and which he used to call type-B physicalism — ‘weak illusionism’ (2018, p. 51). By ‘weak illusionism’ he means a view on which phenomenal experience exists but some of our intuitions about it are illusions. As he puts it: ‘For example, dualist and primitivist intuitions (consciousness is primitive and non-physical) will be incorrect on this model, as will explanatory intuitions (consciousness cannot be physically explained)’ (*ibid.*, p. 51). The illusion at play here is dualism itself. But to label a physicalist view — just because of its denial of dualism — a form of illusionism is a bit tendentious. It is a bit like calling compatibilism about free will ‘free will illusionism’.

As I have argued, explanation is a concept-dependent affair. So whether an explanation needs to feature ‘consciousness’, and so whether or not it is ‘independent’ of consciousness, depends on how we formulate the explanandum. If our beliefs about consciousness are thought about in common-sense terms, we need to explain them in terms of ‘acquaintance’ and its cognates, as I have argued in the previous sections. That reading would render premise 1 false. If, on the other hand, beliefs about consciousness are thought about in physical/functional terms, premise 2 appears to be false, both on the assumption of physicalism, and reasonable versions of dualism. I will only allude here to the physicalist case. Just because there is an explanation of the problem intuitions, stated in physical/functional terms in similarly physical/functional terms, doesn’t say anything about whether consciousness exists, or whether beliefs about it are justified. If type-B physicalism is right, there is an explanatory gap between physical/functional descriptions and phenomenal descriptions, even though consciousness, acquaintance, etc. are purely physical. Premise 2 seems to be begging the question against type-B physicalism as, on type-B physicalism, there *is* a correct explanation of our beliefs about consciousness that is independent of consciousness, even while our beliefs in consciousness are perfectly justified.

There are other arguments for illusionism. Keith Frankish (2016), for example, cites the explanatory gap as reason to deny the existence of phenomenal experience:

If a property resists explanation in physical terms... then the simplest explanation is that it is illusory. In this light, considerations usually cited in support of a radical approach to consciousness, such as the existence of an explanatory gap... afford equal or greater support for illusionism. (*ibid.*, p. 16)

Since, as I have just laid it out, type-B physicalism has a non-illusionist answer to the explanatory gap problem, I do not think this kind of argument is successful either. Instead, in the rest of this paper I will argue that (strong) illusionism is not just a mistaken view, one among many mistaken views regarding the mind–body problem, but one that is intertwined with a general trend to turn away from subjectivity. The fact of its ascendance and its increasing academic and popular visibility is a worrisome sign. Analytic philosophy of mind — for a number of good reasons — has traditionally refrained from cultural criticism. Nevertheless, I think it is instructive to put illusionism in this broader perspective.

Illusionism is the wildly implausible view that there is no conscious experience. It is a view whose apparent absurdity its adherents try in various ways to explain away. Nevertheless, Chalmers appears to be somewhat susceptible to its allures: he confesses that, while he is not an illusionist, *if* he were a physicalist, he would adopt the view:

Dialectically, the illusionist side is much more interesting than the realist side. Looking at the dialectic abstractly, it is easy to sympathize with the illusionist's debunking against the realist's foot-stamping. Still, reflecting on all the data, I think that the realist's side is the right one. (2018, p. 55)

As far as I can see, the allure of illusionism consists in two key features of it. One is its very counter-intuitiveness. Some find the idea of a giant hoax behind our most common-sense view of ourselves exciting. But I think its most important attraction is its adherence to a hardnosed, thoroughgoing scientific outlook. Illusionism appears to uphold the banner of science and reason against the prejudices of common sense.

Granted, science and philosophy can and have gone against deeply held common-sense views. Obvious examples concern the nature of physical objects (containing mostly empty space), the nature of causation (not an inner, unobservable force), or, more controversially, the nature of the self (not a mental substance) and free will (not incompatible with determinism). But the case of experience is not like that. In the case of experience, the pressure that science and philosophy can bring to bear is nowhere near strong enough to justify doubt in one's own experience. There are no scientific discoveries incompatible with the existence of conscious experience; and there are no knock-down philosophical arguments against it, much less a demonstrable incoherence in our concept of it. Consequently, it is not some overwhelming scientific or philosophical reason, accepted on balance and reluctantly in the face of common sense, that draws its adherents to it. Instead, it is rooted in a bias towards scientific and philosophical theorizing and a near-total discounting of the authority accorded to introspective awareness of experience.

I take it that anyone who has seriously considered the nature of subjective experience through introspection will be unable to doubt its existence. But it is possible to turn away from and disregard introspection altogether. Illusionism perhaps sounds plausible, or at least conceivable, from the third-person, scientific perspective on thought, mental representation, and the like. From this point of view, it is possible to argue, as, for example, Keith Frankish (2016) does, that

organisms have no introspective way of checking the accuracy of their introspective representations, and so they cannot rule out the possibility that these representations are non-veridical.

From the scientific perspective, experience appears to be an anomaly; something that is strangely eluding its reach. No matter our metaphysics, the relationship between subjective experience and brain processes remains inexplicable. The explanatory gap doesn't only arise for the physicalist. Even the dualist, who declares experience a non-physical phenomenon, cannot explain how subjectivity fits into an objective world describable by science.⁸ If subjectivity itself is an objective phenomenon, it is hard to explain why it can only be properly grasped subjectively.⁹

This mystery makes experience suspect. So, instead of trying to reconcile the objective and subjective perspectives, scientifically inclined philosophers have come to deny the existence of experience altogether. But, by doing so, they distance themselves from important aspects of human life. Science is no substitute for introspection and self-knowledge. Even if physicalism were true, and so even if we could in principle investigate experience from a scientific perspective as well, it would still be the case that whole areas of human life can only be usefully approached in a subjective, humanistic way.¹⁰ These are areas that are of the greatest importance to human beings, involving matters of value, meaning, and choice.

Suppose there was a super-intelligent organism — let's call her *Zombie Mary*, with a nod to Frank Jackson's (1982) superscientist *Mary* — that lacked any feeling or experience, a creature of pure thought. *Zombie Mary* could know a tremendous amount about humans in biological, neuronal, and information processing terms — even though she lacked the introspective understanding normal humans have of their subjective reality. Nevertheless she would know nothing of value, meaning, and human significance. *Zombie Mary*

⁸ For this point, see Nagel (1979): 'The broader issue between personal and impersonal, or subjective and objective, arises also for a dualist theory of mind. The question of how one can include in the objective world a mental substance having subjective properties is as acute as the question how a physical substance can have subjective properties' (p. 201).

⁹ Panpsychism might provide the most intelligible picture of the mind-brain relationship, but panpsychism is riddled with unsolvable explanatory riddles of its own concerning how proto-conscious states combine to produce full-blown consciousness.

¹⁰ For a discussion, see Ismael (2017).

would not only lack an understanding of pain, but also of the badness of pain. Nothing could be beautiful, or attractive, or horrifying to her; and a case could be made that even moral notions would be incomprehensible to her. If illusionism is right, we are like such a creature (minus the superhuman intelligence).

Luckily, humans are not zombies. We do have experience and value is present in our lives whether we pay attention or not. But if someone thought that experience is an illusion they might not be so inclined to explore the subjective perspective. So, the problem with illusionism is not just that it is false. The problem is that it plays into a modern tendency to turn away from experience. For the most part, this happens in our culture not through science-worship but simply as a result of the speed and overabundance of sensory stimulation, and the lack of encouragement to have an inner life. Illusionism is not the cause of any of this, but it fits in and even provides ideological cover for it.

Acknowledgments

I am grateful for very helpful comments from David Chalmers, François Kammerer, and Barry Loewer.

References

- Baars, B. (1988) *A Cognitive Theory of Consciousness*, Cambridge: Cambridge University Press.
- Balog, K. (1999) Conceivability, possibility, and the mind–body problem, *The Philosophical Review*, **108** (4), pp. 497–528.
- Balog, K. (2012a) Acquaintance and the mind–body problem, in Hill, C. & Gozzano, S. (eds.) *New Perspectives on Type Identity: The Mental and the Physical*, pp. 16–43, Cambridge: Cambridge University Press.
- Balog, K. (2012b) In defense of the phenomenal concept strategy, *Philosophy and Phenomenological Research*, **84** (1), pp. 1–23.
- Block, N. (2006) Max Black’s objection to mind–body identity, in Zimmerman, D. (ed.) *Oxford Studies in Metaphysics*, II, pp. 3–78, Oxford: Oxford University Press.
- Chalmers, D.J. (2003) The content and epistemology of phenomenal belief, in Smith, Q. & Jovic, A. (eds.) *Consciousness: New Philosophical Perspectives*, pp. 220–273, Oxford: Oxford University Press.
- Chalmers, D.J. (2007) Phenomenal concepts and the explanatory gap, in Alter, T. & Walter, S. (eds.) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, Oxford: Oxford University Press.
- Chalmers, D.J. (2018) The meta-problem of consciousness, *Journal of Consciousness Studies*, **25** (9–10), pp. 6–61.

- Frankish, K. (2016) Illusionism as a theory of consciousness, *Journal of Consciousness Studies*, **23** (11–12). Reprinted in Frankish, K. (ed.) (2017) *Illusionism as a Theory of Consciousness*, Exeter: Imprint Academic.
- Ismael, J. (2017) Why (study) the humanities?, in Grimm, S. (ed.) *Making Sense of the World: New Essays on the Philosophy of Understanding*, Oxford: Oxford University Press.
- Jackson, F. (1982) Epiphenomenal qualia, *Philosophical Quarterly*, **32**, pp. 127–136.
- Loar, B. (1990) Phenomenal states, *Philosophical Perspectives*, **4**, pp. 81–108.
- Nagel, T. (1979) *Mortal Questions*, Cambridge: Cambridge University Press.
- Papineau, D. (2002) *Thinking about Consciousness*, Oxford: Oxford University Press.
- Rosenthal, D. (2002) Explaining consciousness, in Chalmers, D.J. (ed.) *Philosophy of Mind: Classical and Contemporary Readings*, New York: Oxford University Press.
- Tononi, G. (2007) Integrated information theory, in Velmans, M. & Schneider, S. (eds.) *The Blackwell Companion to Consciousness*, Oxford: Blackwell.